

Software Citation: State of the practice, challenges, solutions

Stephan Druskat

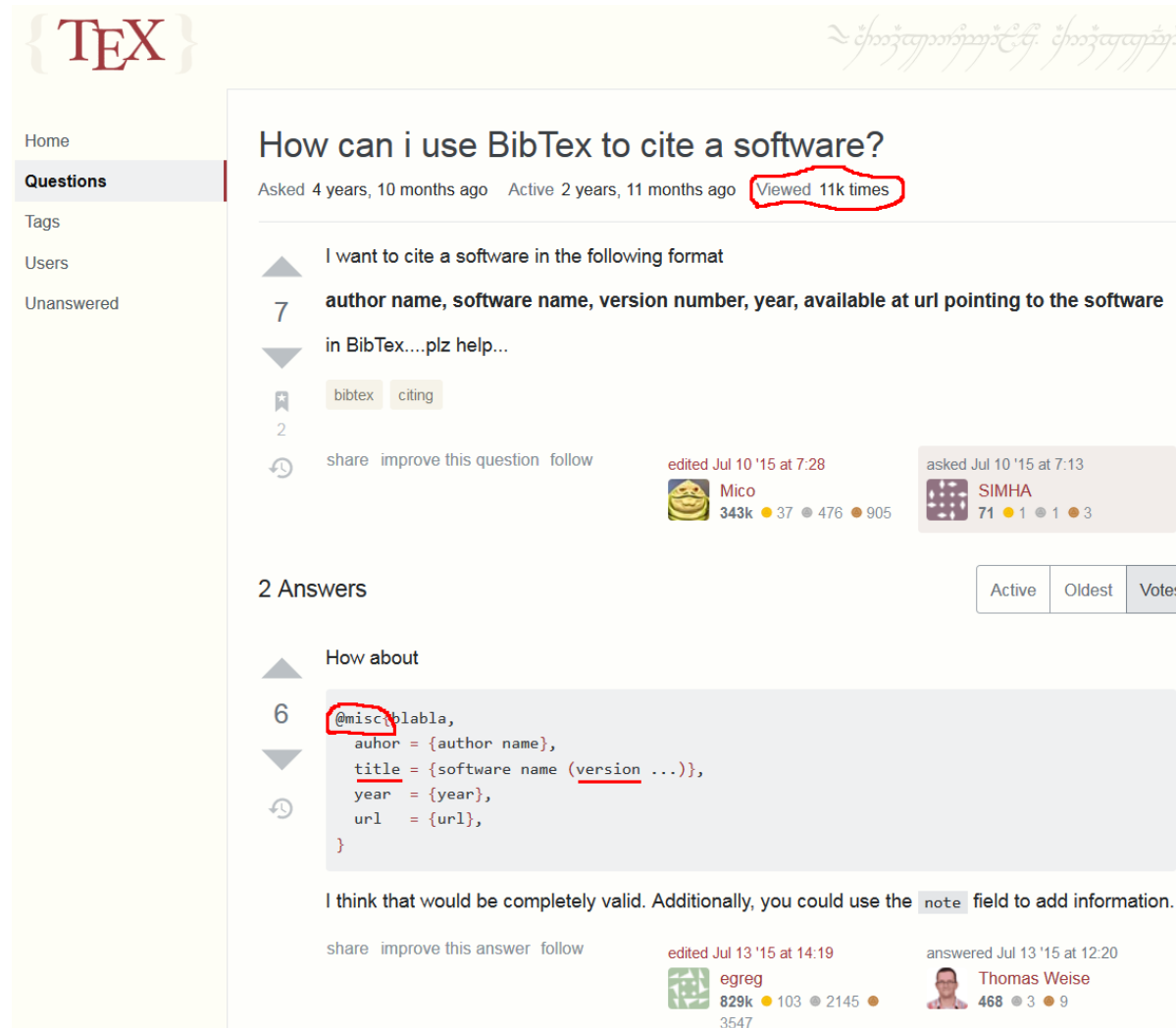
German Aerospace Center (DLR), Institute for Software Technology &
Humboldt-Universität zu Berlin, Dept. of Computer Science &
de-RSE – Society for Research Software

Berliner Bibliothekswissenschaftliches Kolloquium
2020-06-02

A large, curved satellite image of the Earth's surface, showing blue oceans, white clouds, and green landmasses, primarily focusing on Europe and Africa.

Knowledge for Tomorrow

Software citation: A problem?



The screenshot shows a Stack Overflow question page. The question is titled "How can i use BibTeX to cite a software?" and has been viewed 11k times. The asker, Mico, wants to cite software in a specific format: "author name, software name, version number, year, available at url pointing to the software". The question has two answers. The first answer, by @misc:blabla, provides a BibTeX entry template:

```
@misc{blabla,
  author = {author name},
  title = {software name (version ...)},
  year = {year},
  url = {url},
}
```

 The second answer, by Thomas Weise, suggests using the "note" field for additional information. The page also shows a sidebar with navigation links like Home, Questions, Tags, Users, and Unanswered.

How would you cite
<https://github.com/RoboMaster/RoboRTS?>

For https://github.com/luogantt/xgboost_new,
do you cite this?

```
@inproceedings{Chen:2016:XST:2939672.2939785,
  author = {Chen, Tianqi and Guestrin, Carlos},
  title = {{XGBoost}: A Scalable Tree Boosting System},
  booktitle = {Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining},
```

```
...
}
```

https://github.com/luogantt/xgboost_new/blob/master/xgboost-master/CITATION

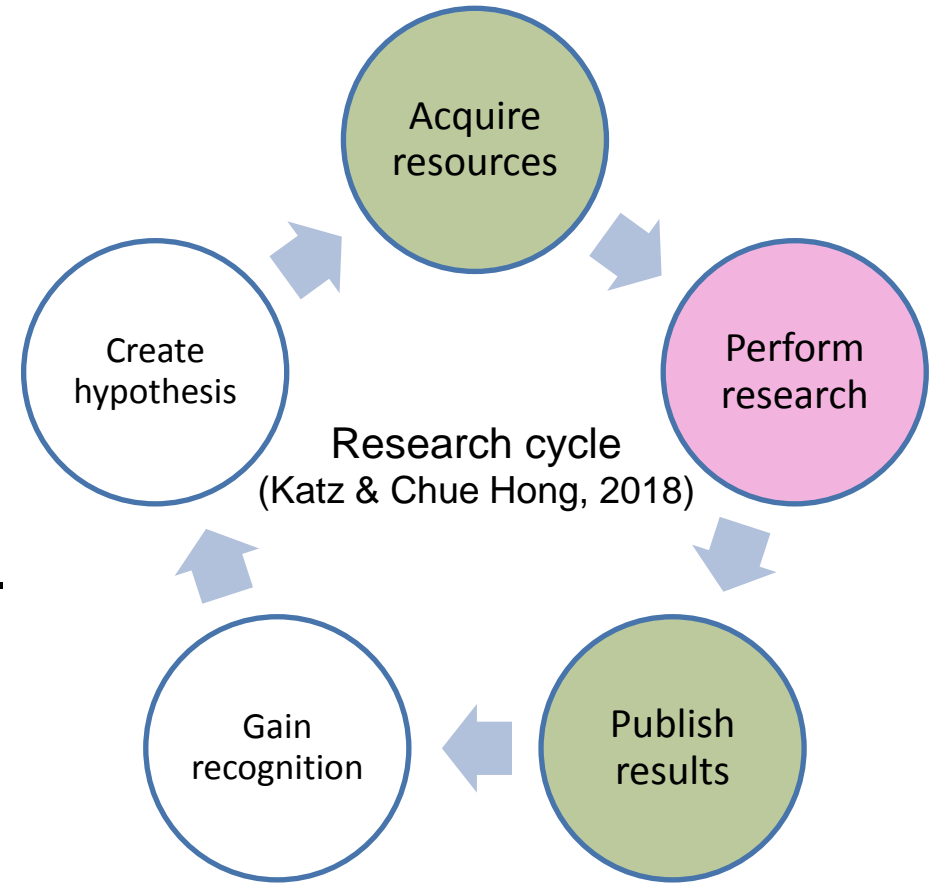
Software Citation: General Challenges

Software is essential to research!

Software **contributes** to the research cycle ...

- 1) ... as *infrastructure*,
- 2) ... as *research product* in its own right.

In research, **contributions** are measured through citation.



Cultural challenge: Software not recognized as research product

Systemic challenge: Current citation system is for papers and books

Overall challenge: Identify software used in research

Software Citation Principles (Smith, Katz, Niemeyer, et al., 2016)

Principles

Meaning: „Software citation ...“

Importance

→ cites software like papers. [*culture, system*]

Credit and attribution

→ facilitates credit and legal attribution. [*culture, system*]

Unique identification

→ uses machine-actionable UUIDs. [*system*]

Persistence

→ lets UUID and metadata persist. [*system*]

Accessibility

→ facilitates access to software, metadata, etc. [*system*]

Specificity

→ facilitates identification of/access to versions. [*system*]



Software Citation Principles: Are We There Yet?

Principles

Example Solutions

Importance

→ Policies, initiatives, guidance, spreading the word

Credit and attribution

→ Software journals, reference lists, policies, creditability

Unique identification

→ DOIs for software

Persistence

→ Archival repositories, Software Heritage

Accessibility

→ Archival repositories

Specificity

→ Versioned DOIs

Also: Different software types! Metadata! Tooling! Stakeholder challenges! Etc. (Katz et al., 2019)

Are we there yet? Not quite! We're making progress, but there are still (tough) challenges, including culture change (which is hard).



Software Citation - What Works Today?

The Citation File Format (CFF) (Druskat et al., 2017)

How to **provide** complete and correct metadata?

Provision of software citation metadata, in machine- and human-readable format, following the Software Citation Principles.

cff-version: 1.1.0

message: If you use this software, cite it with the metadata below.

authors:

- **family-names:** Druskat

- given-names:** Stephan

- orcid:** <https://orcid.org/0000-0003-4925-7248>

title: “My Research Software”

version: 2.0.4

doi: 10.5281/zenodo.1234

date-released: 2017-12-18

citation-file-format.github.io



Software Citation - What Works Today?

biblatex-software style extension (di Cosmo, 2020)

How to **use** complete and correct metadata?

Software entry types for BibL^AT_EX (with Biber backend).

```
@softwareversion {delebecque:hal-02090402-condensed,  
  title = {Scilab},  
  author = {Delebecque, Fran{\c c}ois and Gomez, Claude},  
  url = {https://www.scilab.org/},  
  year = {1994},  
  month = jan,  
  license = {Scilab license},  
  version = {1.1},  
  repository= {https://github.com/scilab/scilab},  
  abstract = {Software for Numerical Computation.}  
}
```

ctan.org/pkg/biblatex-software

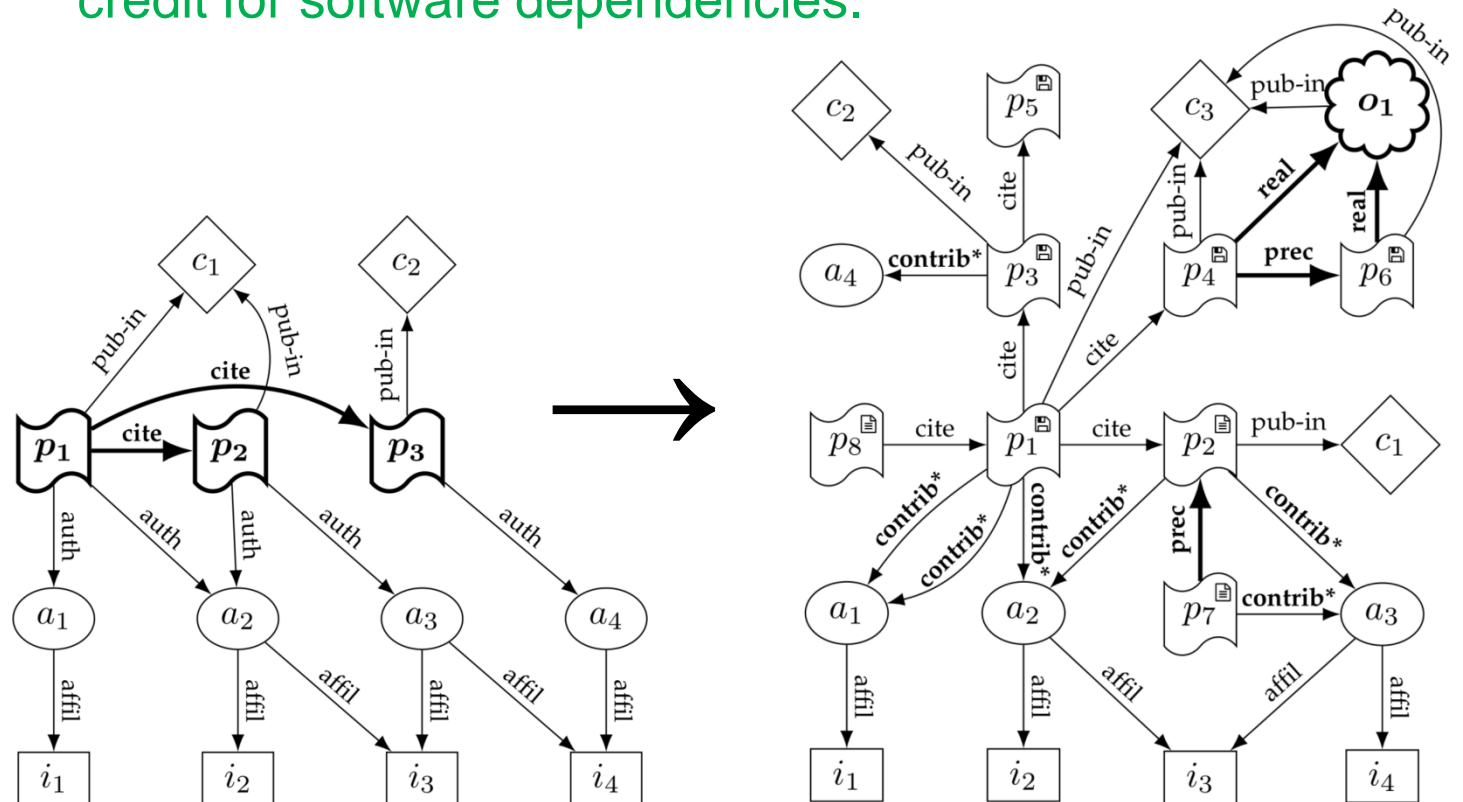


Impact of dependencies in res. software for citation (Druskat, 2020 & fortchoming)

Extend the research citation graph to record transitive credit for software dependencies.

Credit maps!

Credit metric:
B (A): .2, C (B): .3 -> **C (A): .06**



References

- [1] D. S. Katz and N. P. Chue Hong, “Software Citation in Theory and Practice,” in *Mathematical Software – ICMS 2018*, 2018, pp. 289–296.
- [2] A. M. Smith, D. S. Katz, K. E. Niemeyer, and FORCE11 Software Citation Working Group, “Software citation principles,” *PeerJ Comput. Sci.*, vol. 2, no. e86, 2016 [Online]. Available: <https://doi.org/10.7717/peerj-cs.86>
- [3] D. S. Katz *et al.*, “Software Citation Implementation Challenges,” *arXiv:1905.08674 [cs]*, May 2019 [Online]. Available: <http://arxiv.org/abs/1905.08674>.
- [4] S. Druskat, N. Chue Hong, R. Haines, and J. Baker, “Citation File Format (CFF),” 2017 [Online]. Available: <https://doi.org/10.5281/zenodo.1003149>
- [5] R. Di Cosmo, *biblatex-software*. 2020 [Online]. Available: <https://ctan.org/pkg/biblatex-software>
- [6] S. Druskat, ‘Software and Dependencies in Research Citation Graphs’, *Computing in Science & Engineering*, vol. 22, no. 2, pp. 8–21, Mar. 2020, doi: [10.1109/MCSE.2019.2952840](https://doi.org/10.1109/MCSE.2019.2952840). [Online]. Available: <https://doi.org/10.1109/MCSE.2019.2952840>

Figures

- Research cycle diagram adapted from: Daniel S. Katz, “Software Citation in Theory and Practice,” [Online]. Available: <https://www.slideshare.net/danielskatz/software-citation-in-theory-and-practice>.
- Screenshot tex.stackexchange.com: Content under CC BY-SA 3.0 Unported

Thank you!

@stdruskat

stephan.druskat@dlr.de

<https://orcid.org/0000-0003-4925-7248>

<https://de-rse.org>

